Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

AI-Based Prediction of Heat Transfer Performance in Nuclear Power Plant Steam Generators

Saqlain Abbas
College of Nuclear Science and Technology (Harbin Engineering University) Major, Nuclear Science and Technology Email: saqlainabbas68908@gmail.com

ABSTRACT

The overall heat transfer coefficient (U) of a steam generator in a pressurized water reactor (PWR) is a key parameter for both safe operation and fouling control, as well as for efficient power conversion. Classical correlations work well for the dominant physics, but are less effective when fouling, flow, pressure and temperature are all variables with broad operating ranges. An artificial-intelligence model is developed and tested to predict the performance of the heat transfer area of a steam generator using six commonly measured operating parameters. The thermal-hydraulic simulator is a physics-based, correlation-driven simulation that was run with an uncertainty of 2% multiplicative noise to simulate instrumentation uncertainty to produce 3000 operating states. Five regression models: multiple linear regression (MLR), support vector regression (SVR), random forest (RF), gradient boosting regression (GBR) and an artificial neural network (ANN) were tuned, cross validated and tested on a held out set. The best performance of the nonlinear models compared to the linear baseline was for SVR, which produced the highest accuracy (R^2 = 0.9856, RMSE = 116.1 W m-2 K-1, MAPE = 1.73%, with 98.3% of the predictions within ±5% of the reference value, and a 43.5% reduction in RMSE from the linear baseline). The most important contributors to U were fouling resistance (by far), primary inlet temperature, and primary mass flux, as indicated by permutation importance. Robustness tests showed that input perturbations were tolerated well by the tree-based ensembles, but not so well by the SVR and ANN. The findings indicate the potential of using machine learning surrogates that can be easily understood and validated for use as decision support tools in the performance monitoring of steam generators, but highlighted the importance of validating plant data prior to using these surrogates.

Keywords: Steam Generator, Heat Transfer Coefficient, Machine Learning, Fouling, Support Vector Regression, Nuclear Power Plant, Explainable AI

Introduction

Today, nuclear power provides a significant portion of low-carbon electricity globally and the reliability of thermal-hydraulic components plays a pivotal role in the safety and economic operation of nuclear power plants (International Atomic Energy Agency [IAEA], 2020). The steam generator (SG) is the connecting link between the primary coolant circuit and secondary power-conversion circuit in pressurized water reactors (PWRs). Thousands of thin-walled U-tubes circulate hot reactor coolant which

Online ISSN 3007-3197
Print ISSN 3007-3189

extracts the heat by boiling the feedwater on the secondary side to provide the steam to the turbine (Todreas & Kazimi, 2021). Reduced heat transfer capability in this component will result in a decrease in thermal efficiency, a decrease in operating margins and a possible reduction in power levels or unplanned outages.

A number of factors have a strong coupling and control the heat transfer characteristics of an SG, such as the primary-side flow rate and temperature, secondary-side pressure and circulation, feedwater conditions, tube-wall properties, and deposits that have accumulated on the tube surfaces over time. These effects are conventionally modeled empirically and semiempirically, e.g., the Dittus–Boelter relation in single-phase forced convection (Dittus & Boelter, 1985), the Jens–Lottes relation in subcooled and saturated nucleate boiling (Jens & Lottes, 1951), and the Chen superposition approach for convective boiling (Chen, 1966). These tools are vital to design, in addition to other standard heat-exchanger rating methods such as those by Kakaç et al. (2020). Each of these correlations, however, is only valid within a limited parameter range, and the simultaneous variation of more than one variable causes the combination of all the correlations into an overall correlation to lose accuracy, such as when fouling occurs at different rates as the power is moved at different points in the tube bundle or when multiple variables are simultaneously varied during a power maneuver.

A complementary approach for modelling such multivariate and nonlinear relationships is recent machine learning (ML). It has been widely adopted in the field of nuclear science and engineering (NSE) as extensively covered in Gomez-Fernandez et al., 2020; Huang et al., 2023), and, specifically, safety assessment, reliability and explainability (Ayodeji et al., 2022). The heat transfer community has also shown that, for boiling, condensation, and nanofluid flows, data-driven models can match or go beyond the performance of traditional correlations (Baghban et al., 2019; Hughes et al., 2021; Kwon et al., 2020; Qiu et al., 2020; Zhou et al., 2020). Additional studies of critical heat flux (Zhao et al., 2020) and heat-exchanger fouling (Sundar et al., 2020) also suggest that ML can be used to model phenomena that are hard to model analytically. Yet, very few studies consider the overall heat transfer coefficient of a PWR steam generator as a function of operating conditions and fouling, and even fewer incorporate accuracy benchmarking, interpretability and robustness analysis all of which must be performed in order to be accepted in a safety regulated industry (Rudin, 2019; Zio, 2022).

In this paper, we make four contributions. First of all it provides a simple, repeatable procedure for forecasting heat transfer performance of steam generator with ML. Second, it conducts a comparison of five regression families given the same data, tuning, and validation settings. Third, it measures the relative influence of operating variables and fouling by model-agnostic permutation importance measures according to Fisher et al., 2019. Fourth, it tests the robustness of each model against measurement noise, which is a factor that is not often studied but is critical for the deployment of plants (Abdar et al., 2021). The rest of the paper is organized as follows: The second section is a review of the literature, and the questions and hypotheses are presented. The methodology is described in Section 3. The results are presented in Section 4, implications discussed in Section 5, limitations and future work in Section 6, and the conclusions in Section 7.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Literature Review and Research Gap

Conventional Modeling of Steam Generator Heat Transfer

Steam generator thermal performance is conventionally analyzed through the thermal-resistance network that links the primary coolant, tube wall, deposit layer, and boiling secondary fluid (Todreas & Kazimi, 2021). The approach is physically transparent and is embedded in system codes and design handbooks (Kakaç et al., 2020). Its accuracy, however, depends on the validity range of every constituent correlation and on the assumed fouling resistance, which is rarely measured directly and typically inferred indirectly from plant performance trends. High-fidelity computational fluid dynamics can resolve local phenomena but at a computational cost that is incompatible with online monitoring, and even system-level thermal-hydraulic codes require mesh and model choices that introduce discretization-related errors. Data-driven strategies have therefore been proposed to estimate and correct such errors (Bao et al., 2019; Hanna et al., 2020), illustrating the growing role of ML as a complement rather than a replacement for physics-based modeling.

Machine Learning for Heat Transfer Prediction

A large body of work has applied ML to convective and phase-change heat transfer. Artificial neural networks trained on consolidated experimental databases have predicted saturated flow boiling heat transfer coefficients and pressure drop in mini- and micro-channels with accuracy superior to classical correlations (Qiu et al., 2020, 2021), and comparable approaches have been used for flow condensation (Zhou et al., 2020) and for boiling and condensation in finned mini-channels (Zhu et al., 2021). Sensitivity analyses of ML models for CNT/water nanofluid flows in coils highlighted the capacity of data-driven methods to rank the influence of operating variables (Baghban et al., 2019). Deep learning has also been used to infer boiling heat flux from images and data (Liu et al., 2018; Suh et al., 2021). Review articles conclude that ML is most valuable when the underlying physics is partly understood but difficult to close analytically, and when sufficient high-quality data exist (Brunton et al., 2020; Hughes et al., 2021; Kwon et al., 2020). Most of these studies, however, target local or channel-scale quantities, not the integrated, component-level coefficient that matters for plant operators.

Artificial Intelligence in Nuclear Power Plant Systems

In the nuclear field, ML has been applied in the following areas: surrogate modeling of expensive simulations (Radaideh & Kozlowski, 2020), forecasting of accident transients (Radaideh et al., 2020), fault diagnosis (Saeed et al., 2020), and autonomous reactor control architectures (Lin et al., 2021). AI applications in the nuclear field have been explored in reactor transients (Prantikos et al., 2023) and the IAEA (2022) has emphasized the opportunities and governance requirements of AI in nuclear applications. Other concepts such as prognostics and health management (Zio, 2022) and digital twins (Rasheed et al., 2020; Tao et al., 2019) look towards continuous monitoring of components by means of the use of models, for example, steam generators. These studies share the same core message: models are not enough – they need to be explainable and uncertainty-aware, and they need to be robust against

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

measurement errors if they are to be useful for safety-related decisions (Abdar et al., 2021; Arrieta et al., 2020; Rudin, 2019).

Algorithm Selection, Interpretability, and Physics Integration

The selection of an algorithm for the tabular engineering data problem is still being debated. Tree-based ensembles have also recently been found to be very competitive on medium-sized tabular datasets (Grinsztajn et al., 2022; Shwartz-Ziv & Armon, 2022). In the meantime, random forests (Breiman, 2001) and gradient boosting (Friedman, 2001; Chen & Guestrin, 2016) are standard baselines, and the others (kernel methods and shallow neural networks) are good for smooth, low-noise functional relationships. Hyperparameter tuning strategies can greatly affect the ranking results (Probst et al., 2019) and have to be used in a uniform way. Permutation-based and Shapley-based techniques offer model-agnostic explanations and tree-specific explanations, respectively (Fisher et al., 2019; Lundberg et al., 2020). In the meantime, scientific machine learning aims to integrate governing equations or domain knowledge into data-driven models (Cuomo et al., 2022; Karniadakis et al., 2021; Raissi et al., 2019; Willard et al., 2022), a direction which is to the point for nuclear components, where extrapolation outside the training domain can be critical.

Research Gap, Research Questions, and Hypotheses

From the literature surveyed three gaps have been identified. Rather, the studies performed to evaluate heat transfer in a PWR steam generator typically do not explicitly consider fouling as an input and usually do not focus on the overall COT of the steam generator. Second, it is often only the accuracy that is reported in comparative studies, without isolating out which variables are driving the predictions. Third, model robustness to measurement noise (affordable with plant instrumentation) is rarely quantified. The gaps addressed in this study are addressed by the following research questions (RQ) and testing the associated hypotheses (H).

RQ1. To what extent can ML models predict the overall heat transfer coefficient of a PWR steam generator from routinely monitored operating variables, and how do they compare with a linear baseline?

H1. At least one nonlinear ML model will achieve a test-set coefficient of determination of $R^2 \geq 0.98$ and will reduce the root-mean-square error (RMSE) by at least 40% relative to multiple linear regression.

RQ2. Which operating and degradation variables exert the strongest influence on the predicted heat transfer coefficient, and does tube fouling dominate the response?

H2. Fouling resistance will be the single most influential input, with a permutation importance exceeding the combined importance of all remaining variables.

Methodology

Steam Generator Heat Transfer Formulation

The study considers a vertical U-tube PWR steam generator with thin-walled Alloy 690-type tubes. The overall heat transfer coefficient based on the outer tube surface is expressed as a series of thermal resistances (Kakaç et al., 2020; Todreas & Kazimi, 2021), neglecting curvature effects owing to the small wall-to-diameter ratio:

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

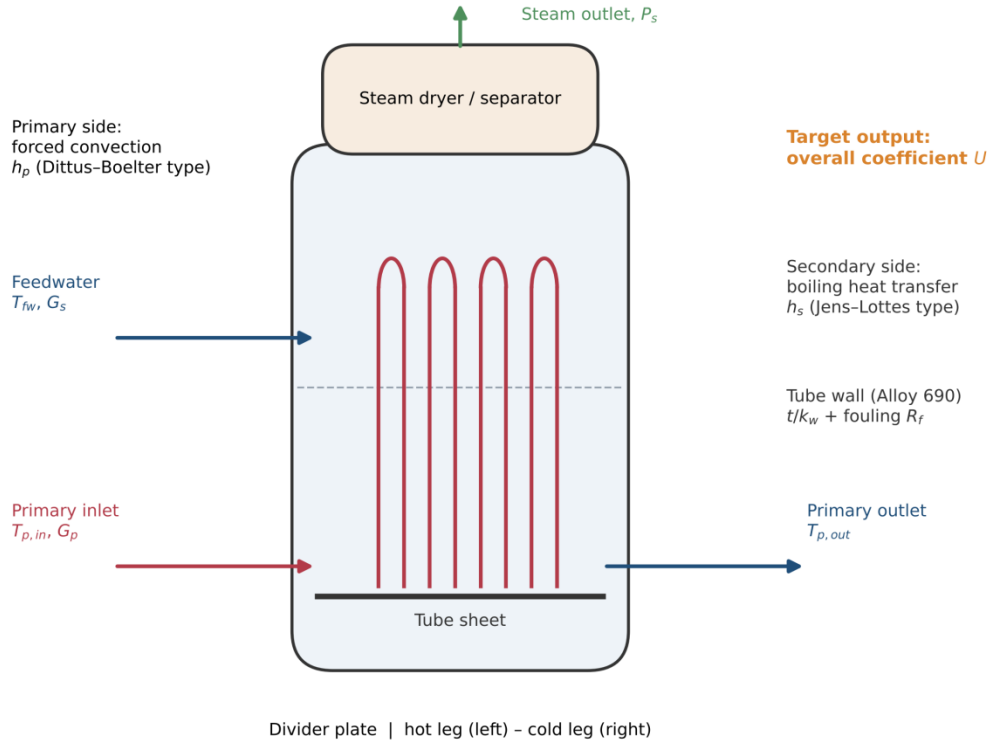
$$1/U = 1/h_p + t_w/k_w + R_f + 1/h_s \quad (1)$$

where h_p and h_s are the primary-side and secondary-side heat transfer coefficients, t_w is the wall thickness, k_w is the wall thermal conductivity, and R_f is the lumped fouling resistance. Figure 1 depicts the system configuration and the variables considered. The primary-side coefficient was evaluated with the Dittus–Boelter correlation (Dittus & Boelter, 1985):

$$Nu = 0.023\, Re^{0.8}\, Pr^{0.4}, \quad h_p = Nu\, k/D_i \quad (2)$$

Figure 1

Schematic of a Pressurized Water Reactor U-Tube Steam Generator and the Variables Considered



with water properties evaluated at the bulk primary temperature, taken as the inlet temperature minus 12 K. Linearized property relations for viscosity, thermal conductivity, and Prandtl number were used over the 300–335 °C range, consistent with the accuracy needed for a surrogate-modeling study. Nucleate boiling on the secondary side was represented by the Jens–Lottes wall-superheat expression (Jens & Lottes, 1951):

$$\Delta T_{sat} = 25\,(q'')^{0.25} \exp(-P/62), \quad h_s = q''/\Delta T_{sat} \quad (3)$$

where q'' is the heat flux in MW m^{-2} and P is the pressure in bar. Following the superposition logic of convective boiling models (Chen, 1966), h_s was multiplied by a circulation factor $(G_s/400)^{0.10}$ and by a feedwater-subcooling correction of the form $1 - 0.0012(\Delta T_{sub} - 60)$, where ΔT_{sub} is the difference between the saturation and feedwater

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

temperatures. Because h_s depends on the heat flux and the heat flux itself depends on U, the system was closed through the relation

$$q'' = U\,(T_b - T_{sat}) \quad (4)$$

and solved by fixed-point iteration until convergence. The wall was assigned a thickness of 1.27 mm and a conductivity of 15 W m^{-1} K^{-1}, and the tube inner diameter was set to 11.11 mm, which are representative of commercial PWR steam generator tubing (Todreas & Kazimi, 2021). The resulting heat fluxes ranged from approximately 56 to 468 kW m^{-2}, with a mean near 203 kW m^{-2}.

Dataset Generation

Monte Carlo sampling method is used to sample six input variables independently and uniformly within the ranges shown in Table 1, which are: primary inlet temperature (Tp,in), primary mass flux (Gp), secondary pressure (Ps), feedwater temperature (Tfw), secondary mass flux (Gs), and fouling resistance (Rf). These range of variables are usually monitored or estimated in plant operations and encompass the standard envelope documented for PWR steam generators (IAEA, 2020; Todreas & Kazimi, 2021). The reference coefficient U was calculated for 3,000 operating states with the above-mentioned formula, and then multiplied by (1 + ε) with Gaussian noise ε having a standard deviation of 2%, in order to simulate the measurement and modelling uncertainty (Abdar et al., 2021). The nature of the data is intentional and is relevant to the use of the data for training and testing data-driven thermal-hydraulic models (Bao et al., 2019; Liu et al., 2018; Radaideh & Kozlowski, 2020). The implications of its external validity are discussed in Section 6.

There are two design considerations worthy of note. The 2% noise level is chosen to be representative of the typical combined uncertainties due to the process instrumentation and correlation scatter, and the level is also sufficiently high to ensure that the learning task is not trivially determinate, which would lead to the preference of interpolating models. Furthermore, in addition, in an operating plant some of the inputs are partially coupled due to control logic, e.g., change of feedwater temperature and secondary pressure with load. Independent sampling has the advantage of providing a uniform coverage of the design space suitable for a methodological comparison but it also involves the fact that the learned models have not been exposed to the correlation structure of real operational data (Hanna et al., 2020).

Table 1

Input Variables and Target Output Used to Build the Steam Generator Dataset

Variable	Symbol	Unit	Range (min–max)	Sampling / source
Primary inlet temperature	$T_{p,in}$	°C	310–335	Uniform
Primary mass flux	G_p	kg m^{-2} s^{-1}	3,000–4,600	Uniform
Secondary pressure	P_s	MPa	5.0–7.0	Uniform
Feedwater temperature	T_{fw}	°C	200–230	Uniform

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Variable	Symbol	Unit	Range (min–max)	Sampling / source
Secondary mass flux	G_s	kg m^{-2} s^{-1}	250–650	Uniform
Fouling resistance	R_f	10^{-5} m^2 K W^{-1}	0–10	Uniform
Overall heat transfer coefficient (target)	U	W m^{-2} K^{-1}	3,531–7,990	2% noise

Note. N = 3,000 operating states. The target has a mean of 5,324 W m^{-2} K^{-1} and a standard deviation of 988 W m^{-2} K^{-1}.

Machine Learning Models

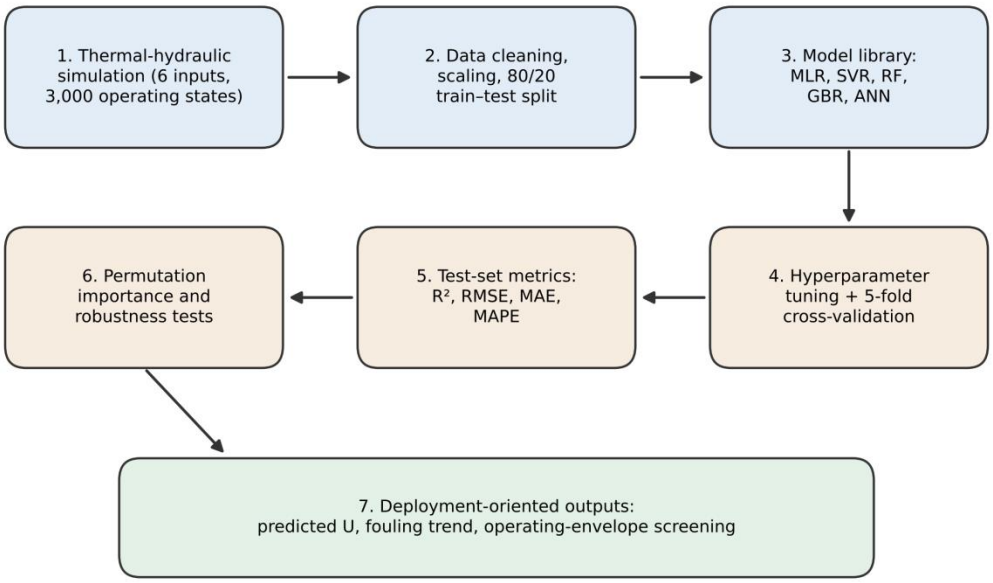
Five regression models were explored. The interpretable baseline was multiple linear regression (MLR). Support vector regression (SVR) with radial basis function (RBF) kernel was chosen due to its capability in handling smooth, low-dimensional problems. Tree-ensemble approaches were widely used and the two most popular approaches were random forest (RF; Breiman, 2001) and gradient boosting regression (GBR; Friedman, 2001), which were among the top performers in recent benchmarks (Grinsztajn et al., 2022; Shwartz-Ziv & Armon, 2022). Lastly, a two-hidden-layer feed-forward artificial neural network (ANN) with 64 and 32 neurons in each layer and rectified linear units (ReLU) as activation functions and Adam optimization (Kingma & Ba, 2015) was added to represent deep learning on tabular data. Python was used to implement all models using the scikit-learn library (Pedregosa et al., 2011). The inputs were transformed to be standardised for MLR, SVR and ANN as these are sensitive to monotonic scaling.Tree-based models were used with raw inputs as they are invariant to monotonic scaling.

Recommended hyperparameter tuning (Probst et al., 2019) was used to select the hyperparameters. For SVR, a grid search with three-fold cross-validation over the penalty parameter C (102, 103, 104) and kernel width γ (0.01, 0.1) with ε = 10 W m-2 K-1 selected C = 10,000 and γ = 0.01. In the implementation of RF, 400 trees were used and 70% of the features were selected at each split. Following the principles of Friedman (2001), the stochastic gradient boosting method, and the regularization philosophy of Chen and Guestrin (2016), GBR used 600 boosting stages, a learning rate of 0.05, a maximum depth of 3, and a subsample fraction of 0.8. The learning rate for the ANN was set to 0.003 and a validation fraction was used to prevent overfitting with an L2 penalty of 10-4. Random seeds were set to ensure the repeatability/control.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Figure 2
Workflow of the AI-Based Heat Transfer Prediction Framework



Validation Strategy and Performance Metrics

The data set was split into two parts, a training set (80%, n = 2400) and an independent test set (20%, n = 600). To estimate the generalization performance, five-fold cross-validation was used on the training set, and the test set was used only once when reporting the final results, thus avoiding the optimistic bias that might result from model selection on the same set. The coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were used to quantify performance. Scale-dependent or percentage-based errors, even when used individually, were not emphasized since they can be misleading when only a single error measure is reported for a model (Chicco et al., 2021); R^2 is more informative than the error magnitudes, RMSE and MAE, in physical units, and relative accuracy, MAPE, is provided. The workflow summarised in Figure 2.

Interpretability and Robustness Analysis

Model-agnostic permutation importance was computed on the test set for the best-performing model, using 30 repeats per variable (Fisher et al., 2019). The importance of a variable is the mean decrease in R^2 when its values are randomly shuffled, and values larger than one are possible because a shuffled variable can drive R^2 below zero. To assess physical consistency, the best model was compared with the noise-free reference solution along two parametric sweeps, one over fouling resistance and one over primary inlet temperature at three fouling levels, with all other inputs fixed at nominal values ($T_{p,in}$ = 322 °C, G_p = 3,800 kg m^{-2} s^{-1}, P_s = 6.3 MPa, T_{fw} = 215 °C, G_s = 450 kg m^{-2} s^{-1}). Robustness was evaluated by perturbing every test input with zero-mean Gaussian multiplicative noise of 2%, 5%, and 10%, without retraining the models,

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

to mimic degraded instrumentation or drift (Abdar et al., 2021).

Results

Predictive Performance of the Models

The performance of the five models is reported in the test-set within Table 2 along with the five-fold cross-validated R^2 of the training data. The other nonlinear models were all considerably better than the MLR baseline model. SVR achieved the highest accuracy (R^2 = 0.9856, RMSE = 116.1 W m-2 K-1, MAE = 92.7 W m-2 K-1, MAPE = 1.73%), followed closely by GBR (R^2 = 0.9846), RF (R^2 = 0.9837), and the ANN (R^2 = 0.9802). The linear baseline reached R^2 = 0.9550 with an RMSE of 205.5 W m-2 K-1. The difference in RMSE between the first four models is relatively small, with a maximum difference of 4.1 W m-2 K-1, which is small relative to the cross-validation variability of the nonlinear models (standard deviations of 0.0006–0.0014 in R^2). The separation between MLR and the nonlinear models, therefore, is clear, but the ranking of the nonlinear models should be taken with a grain of salt.

Table 2

Test-Set Performance and Cross-Validation Results of the Machine Learning Models

Model	R^2	RMSE	MAE	MAPE (%)	CV R^2 (M ± SD)
MLR	0.9550	205.5	163.8	3.09	0.9574 ± 0.0027
SVR	**0.9856**	**116.1**	**92.7**	**1.73**	0.9857 ± 0.0009
RF	0.9837	123.7	97.6	1.83	0.9822 ± 0.0014
GBR	0.9846	120.2	94.8	1.78	0.9848 ± 0.0006
ANN	0.9802	136.3	109.5	2.07	0.9824 ± 0.0009

Note. RMSE and MAE are in W m^{-2} K^{-1}. Test set n = 600; CV = five-fold cross-validation on the training set (n = 2,400). Bold values indicate the best result in each column. MLR = multiple linear regression; SVR = support vector regression; RF = random forest; GBR = gradient boosting regression; ANN = artificial neural network.

The cross validation results showed consistency with the test set results (Table 2) and small standard deviations implied consistent behavior across the data partitions. The RF model appears to have overfitted, with an R^2 of 0.9976 on the training data and 0.9837 on the test data, compared to the GBR model, which has an R^2 of 0.9936 for the training data and 0.9846 for the test data. The typical behaviour of smooth, regularised models fitted to a low-noise function, was that of SVR and ANN, with almost identical training and test scores. The predictive versus the reference coefficients are plotted in figure 3 for the best model in the test set. The points are concentrated along the identity

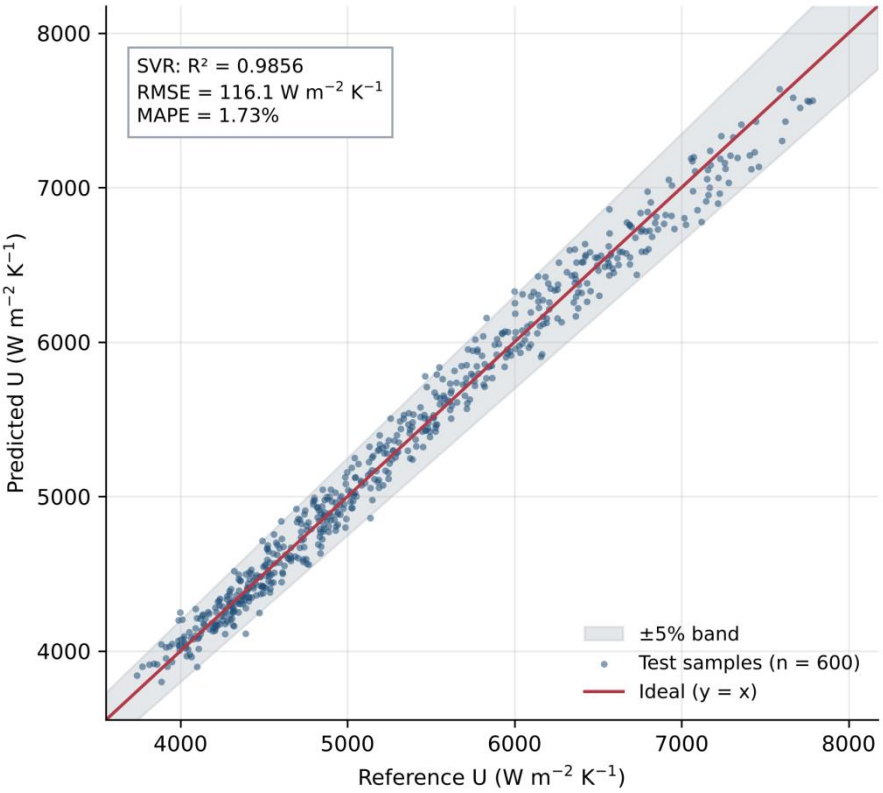
Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

line over the entire range 3,500–8,000 W m-2 K-1: 98.3% of predictions fall within the band ±5% of the identity line (79.7% for MLR). The maximum absolute percent error value found for SVR was 6.4%. The errors were also similar across fouling regimes, with the mean absolute percentage error being 1.85%, 1.71%, and 1.63% in the lowest, middle, and highest terciles of Rf, respectively, suggesting no systematic disadvantage to any regime of fouling.

Figure 3

Predicted Versus Reference Overall Heat Transfer Coefficient for the SVR Model on the Test Set



Note. The shaded area denotes the ±5% error band.

Further insight is gained from learning-curve results. When the SVR model was trained on 240, 600, 1,200, and 2,400 samples, test R^2 rose from 0.9817 to 0.9842, 0.9853, and 0.9856, and RMSE decreased from 131.0 to 121.7, 117.7, and 116.1 W m-2 K-1. In this case, a test R^2 in the range of 0.004 of the final value was obtained with only 240 training data and further samples provided little increase in value. As it is highly encouraging for practical applications, in which high quality labeled plant data are scarce and costly to be obtained, it is also indicative of the smoothness of the response surface.

Influence of Operating Variables and Fouling

The permutation importance of the six inputs to the SVR model is shown in

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

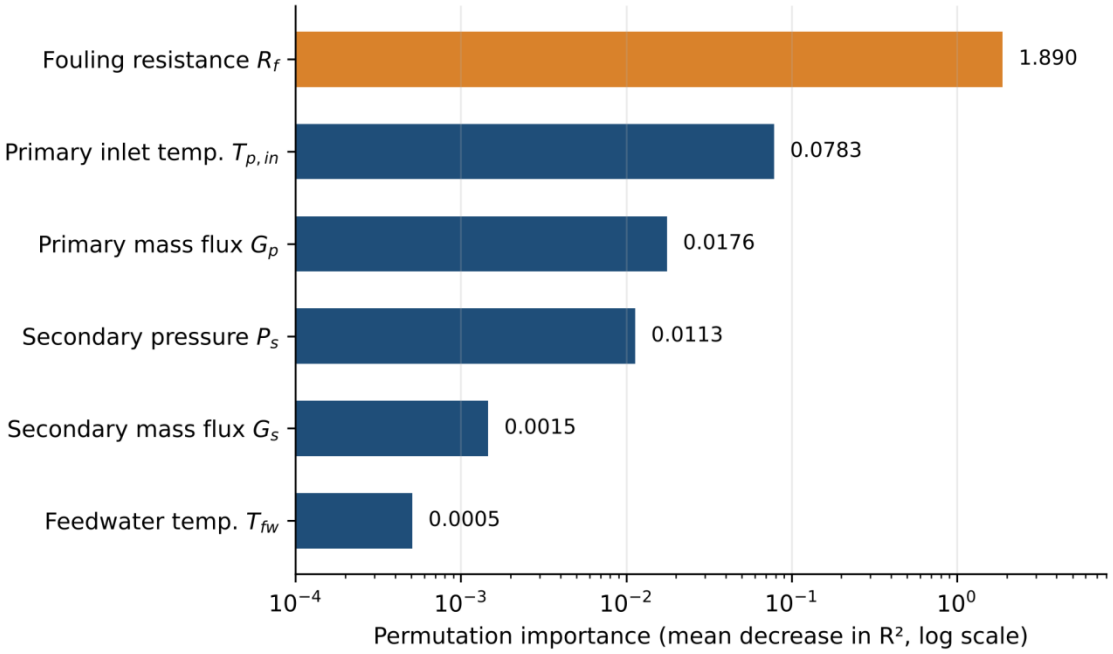
Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

figure 4. The most influential variable was fouling resistance with a mean R2 decrease of 1.890 and over twice that of the next most influential variable (primary inlet temperature with a mean R2 decrease of 0.078). The primary mass flux (0.018) and secondary pressure (0.011) were the next largest contributors, and secondary mass flux (0.0015) and feedwater temperature (0.0005) were the least important. The total importance of all non-fouling variables was around 0.109 or about 5.8% of the fouling importance. This linear relationship between Rf and U was negative and highly significant ($r = -0.954$); the other variables showed only a weak linear relationship ($r < -0.17$ and $r > 0.17$).

Figure 4

Permutation Importance of the Input Variables for the SVR Model



Note. Values are mean decreases in R^2 over 30 repeats on the test set; the horizontal axis is logarithmic.

The ranking is physically sound. In the thermal resistance network, the fouling resistance is of the same order as the sum of the rest of the resistances, and the fouling varies within 0–10 × 10-5 m² K W-1, which results in about 45% change in the overall resistance in the nominal case. The property variations in the primary side and the temperature difference that establishes the heat flux and the boiling coefficient are affected by the primary inlet temperature. The single-phase coefficient is affected by primary mass flux through the Reynolds number, but by a relatively small amount due to the relatively small contribution of the primary resistance to the total. The low influence of feedwater temperature and secondary mass flux is in accordance with low exponents of corresponding correction terms given in the simulator, which means that these parameters are of secondary importance for the integrated coefficient in the range of studied parameters. This is only a provisional conclusion based on the fouling range

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

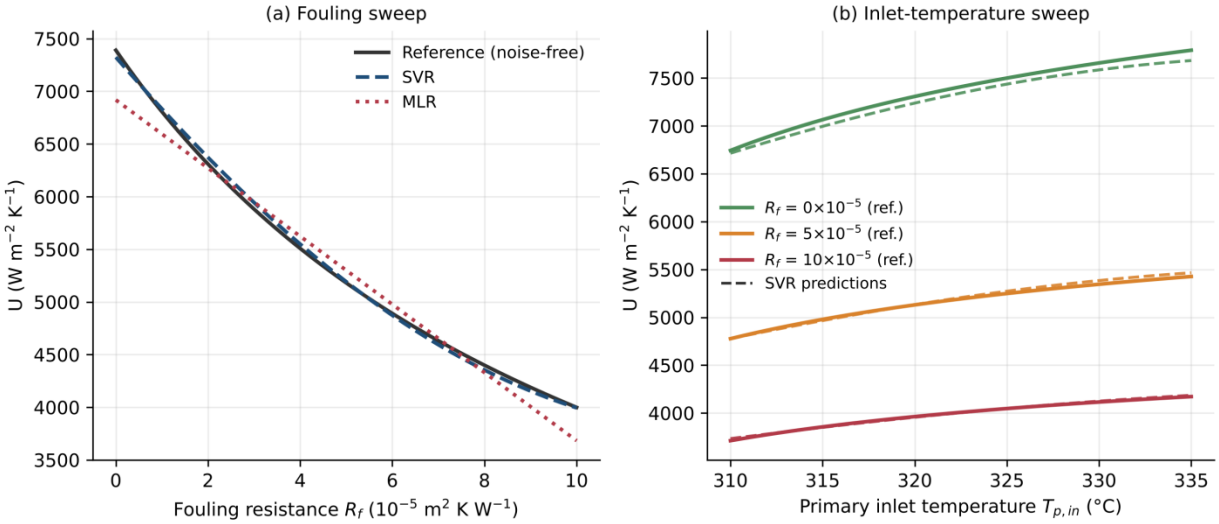
selected in this article and should be reconsidered using plant-specific fouling data.

Physical Consistency Along Parametric Sweeps

Figure 5 shows that the SVR model and the MLR baseline are comparable to the noise-free reference if compared on two parametric sweeps. The profile of the reference coefficient is convex and decelerating, from an initial value of ~7,400 W m-2 K-1 for clean tubes to ~4,000 W m-2 K-1 at Rf = 10 × 10-5 m2 K W-1 (panel (a)). The SVR curve lies almost exactly on the reference over the entire range, while the linear model fails to capture the curve and has a difference of ~6-8% at both extremes of the range. As seen in panel (b), the primary inlet temperature is monotonically increasing with U for all three fouling levels, and the SVR is closely following the trend, underpredicting by about 1–2% at the highest end of the primary inlet temperature range for clean tubes. Importantly, the model matches the correct ordering and spacing of the three fouling curves, indicating that the learned input–output relationship is physically coherent, and not just statistically accurate, as is recommended for ML models intended for engineering decision support (Rudin, 2019; Willard et al., 2022).

Figure 5

Comparison of SVR and MLR With the Reference Solution Along Parametric Sweeps



Note. Solid lines are noise-free reference values; dashed lines are model predictions.

Robustness to Input Perturbations

The results of all models on the effect of input noise are summarized in Table 3. At 2% input noise, every model lost accuracy, but the ordering changed: the tree ensembles (GBR, R^2 = 0.9652; RF, R^2 = 0.9645) overtook SVR (R^2 = 0.9554) and the ANN (R^2 = 0.9500). As noise levels increased, the divergence showed up more significantly. At 5% noise, GBR and RF retained R^2 values of 0.9405 and 0.9390, respectively, while SVR fell to 0.8477, MLR to 0.8199, and the ANN to 0.7551. However, when the noise level was increased to 10%, R^2 of GBR remained at 0.9279 (RMSE = 260.1 W m-2 K-1) while R2 of SVR decreased to 0.6895, R2 of MLR was

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

reduced to 0.4350 and the ANN's R2 was negative (−0.4984), which was worse than using the mean.

Table 3

Robustness of the Models to Multiplicative Gaussian Noise Added to the Test Inputs

Input noise	MLR	SVR	RF	GBR	ANN
0%	0.9550 / 205	0.9856 / 116	0.9837 / 124	0.9846 / 120	0.9802 / 136
2%	0.9309 / 255	0.9554 / 205	0.9645 / 183	0.9652 / 181	0.9500 / 217
5%	0.8199 / 411	0.8477 / 378	0.9390 / 239	0.9405 / 236	0.7551 / 480
10%	0.4350 / 728	0.6895 / 540	0.9252 / 265	0.9279 / 260	−0.4984 / 1,186

Note. Each cell reports R^2 / RMSE (W m^{-2} K^{-1}) on the 600-sample test set; models were not retrained after perturbation.

One possible reason is the way in which both algorithms respond to their inputs which fall outside the envelope of training. Multiplicative perturbations of strongly influential and bounded variables, in particular fouling resistance, with values near zero, move samples to areas of low or no training support. Tree-based models output piecewise-constant values constrained by the training targets, thus restricting the error. In contrast, smooth models are fitted with a surface which is then extrapolated; small changes in the input can lead to large error in the output for these models as has been shown to be the case in Section 4.2. This is in-line of the known capacity of neural networks to extrapolate beyond their training data with unreliable results, and suggests uncertainty-aware and physics-constrained formulations for systems of sensors when reliability is a concern (Abdar et al., 2021; Karniadakis et al., 2021).

One plausible explanation is the way in which each algorithm deals with inputs that fall outside of the training envelope. Samples are pushed to areas with few or no training examples when the variables are multiplicatively perturbed to the extremes, especially the fouling resistance variables, which include values near zero. Tree-based models give piecewise-constant output for such samples and their error is bounded by the training targets, thereby limiting their error. Smooth models, on the other hand, extrapolate their fitted surface; and, as discussed in Section 4.2, they are quite sensitive to shifts in the inputs. This is what has been observed with neural networks in general, and is consistent with the drive for formulations to consider uncertainty of the sensors (Abdar et al., 2021; Karniadakis et al., 2021).

Hypothesis Evaluation

H1 was supported. The best nonlinear model (SVR) reached R^2 = 0.9856, exceeding the 0.98 threshold, and reduced RMSE from 205.5 to 116.1 W m^{-2} K^{-1}, a decrease of 43.5% with respect to MLR, which is above the 40% criterion. All four nonlinear models exceeded the 0.98 threshold (SVR, 0.9856; GBR, 0.9846; RF, 0.9837; ANN, 0.9802), and the corresponding RMSE reductions relative to MLR were 43.5%, 41.5%, 39.8%, and 33.7%, respectively, so that two of the four models also met the 40% error-reduction criterion.

H2 was supported. Fouling resistance had a permutation importance of 1.890,

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

compared with a combined importance of about 0.109 for the five remaining variables. Therefore, the influence of fouling exceeded that of all other inputs together by a factor of roughly 17. It must be emphasized that the magnitude of this dominance is partly conditioned by the fouling range specified in the simulator, whereas the qualitative conclusion, that fouling is the principal degradation-related lever on *U*, is in line with the broader literature on heat-exchanger fouling (Sundar et al., 2020).

Discussion

The results indicate that the overall heat transfer coefficient of a steam generator can be predicted by a small number of six operating variables, less than 2% of the mean absolute percentage error, and using suitable nonlinear regressors. The reduction in RMSE over MLR is 43.5% when using the curvature of the response which proves the usefulness of considering the convex dependency on fouling resistance as illustrated in Figure 5. This is consistent with previous studies indicating ML surrogates can outperform conventional correlations for boiling and condensation heat transfer (Qiu et al., 2020; Zhou et al., 2020) and now for an integrated, component-level performance indicator.

The near-parity of the SVR, GBR, RF and ANN under noise-free test conditions and the switch in accuracy under perturbation is a practical lesson to learn from the results: model selection must not be based solely on the accuracy obtained on the test set. The kernel model and the neural models were good when the response surface was smooth and the inputs were within the training domain, which is in line with the "irregular" theory about the best performance of deep and kernel models on smooth functions, and tree ensembles on irregularities of tabular data (Grinsztajn et al., 2022). The tree ensembles performed most robustly in operational contexts where sensor drift, calibration offsets or missing data were prevalent. A hybrid approach, such as using SVR for nominal monitoring and a tree ensemble as a fallback or verifier, for example, could benefit from both approaches. Another possibility would be to use physics-informed formulations, where the structure of the resistance network in Equation 1 is embedded in the model, which can be helpful for extrapolation (Cuomo et al., 2022; Raissi et al., 2019).

The results of this interpretability analysis have implications for plant management. Due to fouling resistance being the dominant parameter in determining U, there is an inverse relationship between these two that can be inverted in the framework of a digital twin (Rasheed et al., 2020; Tao et al., 2019), allowing the prediction of fouling resistance from regularly measured operating conditions, thereby enabling condition-based cleaning or chemistry intervention as per the concepts of prognostics and health management (Zio, 2022). The learning-curve behavior further indicates that useful surrogates can be trained on limited sets of data, if they are well spread throughout the operating envelope. Lastly, the monotonic and physically-ordered behavior of the parametric sweeps gives a first level of trust, which would need further verification, validation, and uncertainty quantification (Abdar et al., 2021; IAEA, 2022) to obtain regulatory acceptance.

The deployed computational profile of surrogates is also important. All five models predict once trained within a few milliseconds—fast enough to be embedded in

Online ISSN 3007-3197
Print ISSN 3007-3189

plant monitoring software and in the model update loops of digital twins, as compared with days or hours required for a system-code or CFD simulation of a complete steam generator (Brunton et al., 2020; Tao et al., 2019). Retraining the models on the new data from the plants is also cheap, and so the models can be used to track slowly evolving tube conditions. Low-latency, adequate accuracy, and transparency of the importance rankings therefore make these surrogates viable additions to, not alternatives to, physics-based tools and operator judgment (Zio, 2022).

Conclusion

In this study, a novel framework that uses artificial intelligence to forecast the overall heat transfer coefficient of PWR steam generators, based on six operating parameters, was developed and validated. Tuning, cross validation and testing were applied for five regression models on 3,000 operating states. The nonlinear models had R^2 values ranging from 0.980 to 0.986 and the SVR model the lowest error (RMSE = 116.1 W m-2 K-1; MAPE = 1.73%), which was 43.5% lower than the error of the linear baseline. Fouling resistance was the most important parameter in determining the overall heat transfer performance and the optimal model was able to capture the physically anticipated behavior for the parametric sweeps. However, when the input noise was increased, the tree-based ensembles were far more robust than SVR and the neural network. Both hypotheses were confirmed. The results suggest that using ML surrogates in monitoring and decision-support systems for steam generator performance will be relevant in safety-critical applications, if they are validated against plant data, uncertainty quantification is included, and the ML surrogates are integrated into physics-informed frameworks prior to use.

References

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion, 76*, 243–297. https://doi.org/10.1016/j.inffus.2021.05.008

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Ayodeji, A., Amidu, M. A., Olatubosun, S. A., Addad, Y., & Ahmed, H. (2022). Deep learning for safety assessment of nuclear power reactors: Reliability, explainability, and research opportunities. *Progress in Nuclear Energy, 151*, Article 104339. https://doi.org/10.1016/j.pnucene.2022.104339

Baghban, A., Kahani, M., Nazari, M. A., Ahmadi, M. H., & Yan, W.-M. (2019). Sensitivity analysis and application of machine learning methods to predict the heat transfer performance of CNT/water nanofluid flows through coils. *International Journal of Heat and Mass Transfer, 128*, 825–835.

Bao, H., Dinh, N., Lane, J., & Youngblood, R. (2019). A data-driven framework for

Online ISSN 3007-3197
Print ISSN 3007-3189

error estimation and mesh-model optimization in system-level thermal-hydraulic simulation. *Nuclear Engineering and Design, 349*, 27–45.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics, 52*, 477–508. https://doi.org/10.1146/annurev-fluid-010719-060214

Chen, J. C. (1966). Correlation for boiling heat transfer to saturated fluids in convective flow. *Industrial & Engineering Chemistry Process Design and Development, 5*(3), 322–329.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, Article e623. https://doi.org/10.7717/peerj-cs.623

Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F. (2022). Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing, 92*, Article 88. https://doi.org/10.1007/s10915-022-01939-z

Dittus, F. W., & Boelter, L. M. K. (1985). Heat transfer in automobile radiators of the tubular type. *International Communications in Heat and Mass Transfer, 12*(1), 3–22. (Original work published 1930)

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research, 20*(177), 1–81.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gomez-Fernandez, M., Higley, K., Tokuhiro, A., Welter, K., Wong, W.-K., & Yang, H. (2020). Status of research and development of learning-based approaches in nuclear science and engineering: A review. *Nuclear Engineering and Design, 359*, Article 110479. https://doi.org/10.1016/j.nucengdes.2019.110479

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? *Advances in Neural Information Processing Systems, 35*, 507–520.

Hanna, B. N., Dinh, N. T., Youngblood, R. W., & Bolotnov, I. A. (2020). Machine-learning based error prediction approach for coarse-grid computational fluid dynamics (CG-CFD). *Progress in Nuclear Energy, 118*, Article 103140.

Huang, Q., Peng, S., Deng, J., Zeng, H., Zhang, Z., Liu, Y., & Yuan, P. (2023). A review of the application of artificial intelligence to nuclear reactors: Where we are and what's next. *Heliyon, 9*(3), Article e13883.

Online ISSN 3007-3197
Print ISSN 3007-3189

Hughes, M. T., Kini, G., & Garimella, S. (2021). Status, challenges, and potential for machine learning in understanding and applying heat transfer phenomena. *Journal of Heat Transfer, 143*(12), Article 120802.

International Atomic Energy Agency. (2020). *Advances in small modular reactor technology developments: A supplement to IAEA Advanced Reactors Information System (ARIS), 2020 edition*. IAEA.

International Atomic Energy Agency. (2022). *Artificial intelligence for accelerating nuclear applications, science and technology*. IAEA.

Jens, W. H., & Lottes, P. A. (1951). *Analysis of heat transfer, burnout, pressure drop and density data for high-pressure water* (Report No. ANL-4627). Argonne National Laboratory.

Kakaç, S., Liu, H., & Pramuanjaroenkij, A. (2020). *Heat exchangers: Selection, rating, and thermal design* (4th ed.). CRC Press.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics, 3*(6), 422–440. https://doi.org/10.1038/s42254-021-00314-5

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*.

Kwon, B., Ejaz, F., & Hwang, L. K. (2020). Machine learning for heat transfer correlations. *International Communications in Heat and Mass Transfer, 116*, Article 104694.

Lin, L., Athe, P., Rouxelin, P., Avramova, M., Gupta, A., Youngblood, R., Lane, J., & Dinh, N. (2021). Development and assessment of a nearly autonomous management and control system for advanced reactors. *Annals of Nuclear Energy, 150*, Article 107861.

Liu, Y., Dinh, N., Sato, Y., & Niceno, B. (2018). Data-driven modeling for boiling heat transfer: Using deep neural networks and high-fidelity simulation results. *Applied Thermal Engineering, 144*, 305–320.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Prantikos, K., Chatzidakis, S., Tsoukalas, L. H., & Heifetz, A. (2023). Physics-informed neural network with transfer learning (TL-PINN) based on domain similarity measure for prediction of nuclear reactor transients. *Scientific Reports, 13*, Article 16840.

Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. *WIREs Data Mining and Knowledge Discovery, 9*(3), Article e1301.

Online ISSN 3007-3197
Print ISSN 3007-3189

Qiu, Y., Garg, D., Zhou, L., Kharangate, C. R., Kim, S.-M., & Mudawar, I. (2020). An artificial neural network model to predict mini/micro-channels saturated flow boiling heat transfer coefficient based on universal consolidated data. *International Journal of Heat and Mass Transfer, 149*, Article 119211.

Qiu, Y., Garg, D., Kim, S.-M., Mudawar, I., & Kharangate, C. R. (2021). Machine learning algorithms to predict flow boiling pressure drop in mini/micro-channels based on universal consolidated data. *International Journal of Heat and Mass Transfer, 178*, Article 121607.

Radaideh, M. I., & Kozlowski, T. (2020). Surrogate modeling of advanced computer simulations using deep Gaussian processes. *Reliability Engineering & System Safety, 195*, Article 106731.

Radaideh, M. I., Pigg, C., Kozlowski, T., Deng, Y., & Qu, A. (2020). Neural-based time series forecasting of loss of coolant accidents. *Annals of Nuclear Energy, 144*, Article 107557.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics, 378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access, 8*, 21980–22012.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Saeed, H. A., Peng, M.-J., Wang, H., & Zhang, B.-W. (2020). Novel fault diagnosis scheme utilizing deep learning networks. *Progress in Nuclear Energy, 118*, Article 103066.

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion, 81*, 84–90. https://doi.org/10.1016/j.inffus.2021.11.011

Suh, Y., Bostanabad, R., & Won, Y. (2021). Deep learning predicts boiling heat transfer. *Scientific Reports, 11*, Article 5622.

Sundar, S., Rajagopal, M. C., Zhao, H., Kuntumalla, G., Meng, Y., Chang, H. C., Shao, C., Ferreira, P., Miljkovic, N., Sinha, S., & Salapaka, S. (2020). Fouling modeling and prediction approach for heat exchangers using deep learning. *International Journal of Heat and Mass Transfer, 159*, Article 120112.

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics, 15*(4), 2405–2415.

Todreas, N. E., & Kazimi, M. S. (2021). *Nuclear systems: Vol. 1. Thermal hydraulic fundamentals* (3rd ed.). CRC Press.

Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Computing Surveys, 55*(4), Article 66. https://doi.org/10.1145/3514228

Zhao, X., Shirvan, K., Salko, R. K., & Guo, F. (2020). On the prediction of critical heat flux using a physics-informed machine learning-aided framework. *Applied Thermal Engineering, 164*, Article 114540. https://doi.org/10.1016/j.applthermaleng.2019.114540

Online ISSN 3007-3197
Print ISSN 3007-3189

Zhou, L., Garg, D., Qiu, Y., Kim, S.-M., Mudawar, I., & Kharangate, C. R. (2020). Machine learning algorithms to predict flow condensation heat transfer coefficient in mini/micro-channel utilizing universal data. *International Journal of Heat and Mass Transfer, 162*, Article 120351.

Zhu, G., Wen, T., & Zhang, D. (2021). Machine learning based approach for the prediction of flow boiling/condensation heat transfer performance in mini channels with serrated fins. *International Journal of Heat and Mass Transfer, 166*, Article 120783.

Zio, E. (2022). Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice. *Reliability Engineering & System Safety, 218*, Article 108119. https://doi.org/10.1016/j.ress.2021.108119